Poster II-5

A Hybrid Computational/ Experimental Approach to Protein Structure Determination Employing High-Throughput Amide Hydrogen/Deuterium Exchange Mass Spectrometry (DXMS)

Pantazatos, Dennis P.[1], Hilser, Vince[2], Bourne, Phil[3], Baker, David[4], White, Halbert[5], Torrance, James[3], Gessner, Chris[1], Woods, Virgil L. Jr.[1]

[1]Departments of Medicine and [5]Economics, University of California at San Diego, La Jolla, CA, USA; [2]University of Texas Medical Branch, Galveston, TX, USA; [3]San Diego Supercomputing Center, La Jolla, CA, USA; [4]Department of Biochemistry, University of Washington, Seattle, WA, USA

3-D protein structure prediction from primary sequence information remains a holy grail of molecular biology, and success in this effort promises great speed in determination of protein structures. Progress is being made, as evidence from the Critical Assessment of techniques for protein Structure Prediction (CASP) experiments, but that progress has been incremental. We propose a new approach to structure determination that first employs purely computational methods (the Rosetta algorithm) to predict a large number of likely 3-D structures for a protein and then, for each structure prediction generated, calculate (with the COREX algorithm) the anticipated hydrogen exchange rates of each of the protein's peptide amides; its exchange- rate "fingerprint". Simultaneously, the actual exchange- rate fingerprint of the protein is experimentally determined using microgram quantities of the protein, employing high throughput embodiments of peptide amide hydrogen/deuterium exchange lc- mass spectrometry ("DXMS"). Accurate structural predictions are discerned from less accurate ones by comparison of each prediction's calculated rate fingerprint with the experimentally determined rate fingerprint. We term this hybrid approach to structure determination the "DXMS-Rosetta-COREX" filter.

Considerable theoretical and experimental evidence has established that the exchange rates of the peptide amide bond hydrogens within proteins are exquisitely dependent upon protein structure and thermodynamic stability. We (V.W.) have recently developed and integrated a number of enhancements to amide hydrogen/ deuterium exchange- mass spectrometry (which we term DXMS technology) that allow the exchange rates of all of a protein's peptide amide hydrogens (its exchange rate "fingerprint") to be determined in days. Furthermore, we (V.H.) have pioneered the development of computational approaches (the COREX algorithm) capable of reliably predicting amide hydrogen exchange rate fingerprints from actual or presumed protein structure(s). Finally, we (D.B.) have developed the Rosetta algorithm, widely considered one of the more successful *ab-initio* 3-D protein structure predictive methods available. These three methods are combined to form the DXMS-Rosetta-COREX filter as follows:

1. Microgram quantities of the protein are produced and studied by DXMS technology to measure amide hydrogen exchange rates, establishing the protein's true exchange rate fingerprint.
2. Multiple structures (1,000-10,000) are predicted for a target protein using the Rosetta algorithm as well as a variety of other methods.
3. COREX, run on the Blue Horizon supercomputer at the San Diego Supercomputer Center, and on other platforms, is used to construct a virtual hydrogen- exchange- rate fingerprint for each of several proposed structure(s) for a target protein.
4. These calculated fingerprints are compared to the true solution- phase rate fingerprint of the target established by simultaneous DXMS- study of the protein, and the structural predictions with calculated exchange rate fingerprints most closely matching experimentally determined fingerprints identified and ranked.

A description of the DXMS- Rosetta - COREX filter will be presented along with preliminary results obtained with two CASP4 targets showing that, even at this early stage of development, the filter can discern accurate predictions from 80% of less accurate predictions.

This work was supported in part by NIH NCI IMAT Grant CA99835 (V.W.)